Estimating Counterfactual Distributions under Interference

Shiv Shankar SSHANKAR@CS.UMASS.EDU
University of Massachusetts

Ritwik Sinha RISINHA@ADOBE.COM
Adobe Research

Madalina Fiterau MFITERAU@CS.UMASS.EDU
University of Massachusetts

Abstract

Randomized control trials (RCTs) form a key tool for evaluating medical treatments. However, commonly used techniques in RCT based treatment effect estimates suffer from two issues: a) ignoring spillover effects and b) focusing on average (or conditional) average effects. This is partly because evaluating counterfactual distributions is hard; and is further complicated by presence of interference. In this work we propose a new estimator, named RUMI, for estimating distributional quantities like CVaR, QTE from controlled trials under known interference. We provide theoretical justification behind our method and demonstrate its application on synthetic and semi-synthetic experiments.

Keywords: Treatment Effect, Interference, Counterfactual Distribution, RCT

Data and Code Availability

- Data: We conduct synthetic experiments based on random graph simulations following Cortez-Rodriguez et al. (2023). We also evaluate our method on the publicly released samples by Song et al. (2020).
- Code: Code release at `https://github.com/sshivs/rumi`.

Institutional Review Board (IRB) Our research uses publicly available data and/or simulated experiments. This work does not require IRB approval.

1. Introduction

Randomized Control Trials (RCTs), sometimes also known as A/B testing is one of the most effective ways to evaluate interventions, whether it be a new policy (Papadogeorgou et al., 2020) or a medical treatment (Antman et al., 1992), or experimentation in the digital world (Siroker and Koomen, 2015). These allow estimation of the treatment effect by ensuring that treatment (group A) and control (group B) assignments are made independently of other variables, including potentially unknown ones. This is done by randomly assigning units (for example patients enrolled) to one of two candidate vaccines, and the relative response of each of the groups assessed based on a metrics of interest, such as health, infection rates etc. In other cases, interaction data logged from previous deployments can be used for an initial screening of a new algorithm(Kohavi and Longbotham, 2015).

While RCTs are a gold-standard evaluation procedure, there are many scenarios, where critical assumptions required for the validity of such tests are violated. In certain scenarios, there exist underlying mechanisms through which the exposures to some units also affect other units, sometimes known as interference (Hudgens and Halloran, 2008; LeSage and Pace, 2009). We list below some cases where this problem can manifest:

(A) Off-Service Patient Care: Spillover effects can occur when patients are assigned to off-service care, i.e., a ward or department that does not specialize in their condition. The quality of care for both the off-service patients and those in the specialized ward may suffer, as resources like staff attention, equipment, and specialized knowledge become diluted. This creates interdependence between units, where the treatment of one patient indirectly affects the care and outcomes of others, leading to biased estimates of treatment effects.

(B) Resource Constraint: Hospitals often pool and limit certain resources, for efficiency and cost reduction purposes. If only a few patients undergo

specific treatments, the hospital's resources—such as sterile operating rooms, skilled surgeons, and nursing staff—are sufficient to maintain high-quality outcomes. However, as the number of such patients increases, these limited resources become stretched. This leads to interference between patients, where the outcomes of one treatment may be affected by the scarcity of resources for another. For instance, overcrowded operating rooms or fatigued surgeons might result in longer wait times or lower-quality care.

(C) Epidemic Control: Vaccination reduces infection rates not only via the direct effect on vaccinated individuals, but also by interrupting transmission chains. This is a well known phenomenon in epidemiology known as *herd immunity*, and ignoring it leads to underestimation in the treatment effect.

Interference causes the estimated treatment effect to be biased, reflecting not just the treatment on the unit, but other factors. While interference-aware methods for treatment effect estimation exist, they usually focus on the *average effect* (Ogburn et al., 2017; Leung, 2020). However, practical or ethical considerations means that practitioners are also concerned with metrics such as *quantile values* and *performance at a given risk (CVaR)*. These measures are important as the average effect might not reflect the deterioration of outcomes within specific subgroups (e.g., patients requiring more complex surgeries). These quantities need estimation of not just the average but also distribution of outcomes (Altschuler et al., 2019). However, to the best of our knowledge, there does not exist any work on evaluating the counterfactual distribution under interference.

Contributions. We focus on the challenge of estimating the risk-metrics like quantiles, variance, conditional value-at-risk (CVar) etc. and more generally the population level counterfactual distribution in presence of interference among units. We consider the setting where there is an interaction graph among subjects, and interference emanates from 'neighboring' subjects. Compared to prior work, our focus is on estimating the entire outcome distribution instead of average outcomes and treatment effect.

In this work, we propose RUMI a method that allows for estimation of counterfactual distributions (Theorem 1) under interference. We do this by using a functional additive decomposition of the outcome distributions Han et al. (2019). This makes it possible to perform estimation in presence of interference without incurring high variance. Our paper not only establishes the theoretical validity of our estimator but also substantiates its practical efficacy.

2. Preliminaries

2.1. Notation

Let there be n units in the population, $\boldsymbol{Z}$ be the treatment assignment vector of the entire population, e.g., for binary treatments $\boldsymbol{Z} \in \{0,1\}^n$. We use the Neyman potential outcome framework (Neyman, 1923; Rubin, 1974), and let $Y_i(\boldsymbol{z})$ denote the potential outcome of the i-th unit for $\boldsymbol{z} \in \mathcal{Z}$. We have data recorded from an RCT that assigns treatments on the unit i with $Z = 1/0$ corresponds to the treatment and control respectively. Let Y_i be the observed outcome, and X_i be corresponding covariates for unit i. The treatment assignment is done independently at each node i with probability $p_i(Z = Z_i|X_i)$.

We assume that the effect of interference between units can be encoded as a graph. Specifically, we assume that the interference at a device is limited to its immediate neighbours in the graph (i.e. SUTVA holds at the network level). In such a case, the dependence between units can be encoded as a matrix A with $A_{ij} = 1$ only if treatment at unit j can influence outcomes at unit i (e.g. if units j is a friend of unit i or if i and j are used by the same user). We denote by $\mathcal{N}_i = \{j : A_{ij} = 1\}$ the set of *neighbors* of unit i.

The cumulative outcome distribution observed for a treatment allocation policy π is denoted by $F^\pi()$. The sample estimate of the distribution will be denoted by $\hat{F}^\pi()$. Similarly, other sample estimates are denoted by $\hat{}$ mark. For simplicity we will focus on the distribution F^1 which is the outcomes for when all units are allocated to the treatment A. Finally for notational convenience, we focus on only the outcomes corresponding to the treatment i.e. we consider the case of allocating $Z_k = 1$ for all units k. But the method trivially generalizes to any allocation policy π which randomizes between $Z = 0$ and $Z = 1$.

2.2. Related Work

Network interference, also known as interference or spillover effects, occurs when the treatment or exposure of one unit in a network affects the outcomes of other units in the network (Hudgens and Halloran, 2008; Leung, 2020). Various methods have been proposed to address network interference in causal inference, each making different assumptions about the nature of the exposure function (Auerbach and Tabord-

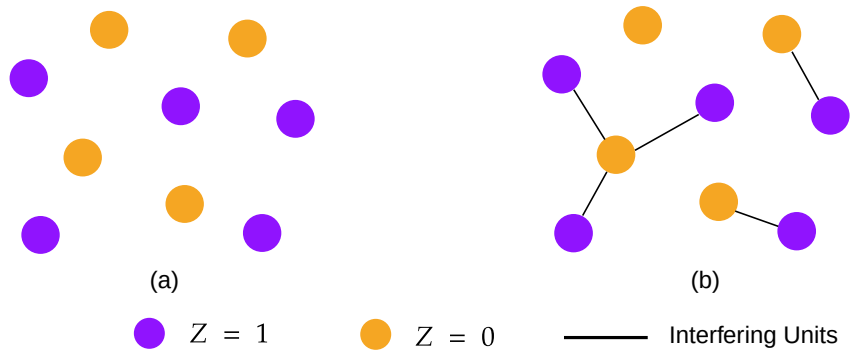


Figure 1: $Z = 1$ denotes to units in the treatment group and $Z = 0$ denotes units in the control group. **(a)** Standard RCT testing where there is no interaction between the treatment and the control units. **(b)** Network interference due to (unknown) interaction between the units.

Meehan, 2021; Li et al., 2021; Viviano, 2020) and the structure of interference neighborhoods (Bargagli-Stoffi et al., 2020; Sussman and Airoldi, 2017; Ugander et al., 2013). Two common assumptions in the literature include a) linearity with respect to a known functional of neighbor treatments (Basse and Airoldi, 2018; Cai et al., 2015; Chin, 2019; Gui et al., 2015; Parker et al., 2017) ; and b) exposure represented as the (weighted) proportion of either the neighboring units that have received treatment (Eckles et al., 2017; Toulis and Kao, 2013), or the count of neighboring units that have undergone treatment (Ugander et al., 2013). Such settings within the context of bipartite interference have been explored by researchers such as Pouget-Abadie et al. (2017), Pouget-Abadie et al. (2019), and Brennan et al. (2022). Furthermore, when the exposure model is known, various cluster randomized designs have been proposed for variance reduction in the estimate (Eckles et al., 2017; Gui et al., 2015; Pouget-Abadie et al., 2019).

Estimating mean outcomes from recorded data is a crucial task for evaluating the effectiveness of different algorithms and policies; and hence has been studied in other areas such as reinforcement learning (RL) (Dudík et al., 2014; Wang et al., 2017) and recommendation systems (Swaminathan et al., 2017; Vlassis et al., 2021). The methods used in these areas are primarily versions of the Horvitz-Thompson estimator (Horvitz and Thompson, 1952).

For comprehensive analysis, relying solely on the expected value is imprudent. Metrics such as Conditional Value at Risk (CVaR) and other measures dependent on the entire reward distribution become essential for evaluating strategies, particularly in high-risk scenarios where the risk or uncertainty associated with different outcomes is a concern (Keramati et al., 2020; Altschuler et al., 2019). Furthermore, assessing the distribution of outcomes is crucial not only for risk assessment but also for gaining insights into the underlying structure of the data. For instance, detecting multimodal or irregular responses to treatment can indicate the presence of unknown groups or clusters within the data (Kennedy et al., 2021).

Counterfactual distribution estimation has also been a subject of considerable research (Robins and Rotnitzky, 2001; Chernozhukov and Hansen, 2005; Westling and Carone, 2020). Most methods in this direction use Horvitz-Thompson-esque estimators (Abadie, 2002; Chernozhukov and Hansen, 2005; Chernozhukov et al., 2013; Díaz, 2017) to estimate the cumulative distribution. Some research has looked at other methods for counterfactual density estimation has been tackled with methods like like kernel smoothing (Robins and Rotnitzky, 2001; Kim et al., 2018) or parametric projection methods (Westling and Carone, 2020; Kennedy et al., 2021). However, all of these methods assume the Stable Unit Treatment Value Assumption (SUTVA), and do not work in the presence of interference (Athey et al., 2018; Sävje et al., 2021).

3. RUMI

Next we present our method called RUMI (Risk and Uncertainity Metrics under Interference), to address the problem of evaluating counterfactual distribution under interference. The principle methodology for

RUMI uses the framework of Abadie (2002); Chernozhukov et al. (2013) who suggest using IPW weighting to estimate the CDF of a counterfactual outcome from observational logs. This estimator follows directly from the definition of the density function and the method of inverse propensity scoring (IPS); sometimes also called importance weighting.

$$F_Y^{\pi}(t) = \Pr(Y \leqslant t) = \mathbb{E}[\mathbb{I}[Y < t]] \overset{a}{=} \mathbb{E}[\mathbb{I}[Y < t]\frac{dF^{\pi}}{dF^{p}}$$

where in (a) we have used the propensity weighting trick to change the original expectation over distribution F^{π} to and expectation over the distribution F^{p}. The function in the expectation is reweighed by the Radon-Nikodym derivative (also known as inverse propensity or importance ratio) between F^{π} and F^{p} [1]. A Monte-Carlo version of this expression can be obtained by replacing the expectation over means drawn from logged data. This estimator for our scenario of network interference is known (Aronow and Middleton, 2013; Hudgens and Halloran, 2008) to reduce to:

$$\begin{aligned}\hat{F}^{1}_{\mathrm{HT}}(t) &= \frac{1}{n}\sum_i \mathbb{I}[Y_i < t] \prod_{j \in \mathcal{N}_i}\left(\frac{z_j}{p}\right) \\ &= \frac{1}{n}\sum_i \mathbb{I}[Y_i < t]\rho^i_{\mathrm{HT}}. \end{aligned} \tag{1}$$

However, using $\hat{F}^{1}_{\mathrm{HT}}$ is practically infeasible because of the *exponentially* high variance due to the product over multiple IPS values. Unfortunately this variance is in the most general setting unavoidable Aronow and Samii (2017). This is partially because without stronger assumptions, the degree of freedom for the model) is exponential in number of neighbours $|\mathcal{N}_i|$.

3.1. Assumptions

Additive Densities: The problem of exponentially high degrees of freedom is well known in treatment effect literature (Toulis and Kao, 2013; Sussman and Airoldi, 2017). It can be tackled, if there is additional structure restricting the set of potential outcome functions (Eckles et al., 2017; Toulis and Kao, 2013; Cortez et al., 2022; Kallus and Uehara, 2019). A common assumption used is the heterogeneous linear interference model (Eckles et al., 2017; Sussman and Airoldi, 2017; Swaminathan et al., 2017), Analogous to such additive outcome models we posit a additive functional density model (Han et al., 2019; Müller and Yao, 2008). This allows the CDF of the outcome variable to be decomposed as a sum of unit-level functions (Assumption 1). that allows *consistent* estimation of the *distribution* under global treatment.

Additive Density

Assumption 1 *The (conditional) cumulative density function (CDF) of the outcomes can be written as a linear combination of unknown unit level latent functions:*

$$F_Y(t) = \sum_{k=1}^{K} \psi_k(Z_k, X, t), \ \forall t \qquad \textbf{(A1)}$$

where $F_Y(t) := \Pr(Y \leqslant t \mid \boldsymbol{Z}, X)$.

We consider this unitwise additive model to explain the core idea of our estimator. In Section 3.3 we consider the generalized setting incorporating interference that happens only when multiple neighbours are treated.

In addition to Assumption 1, we will also posit standard assumptions of network ignorability, positivity and consistency from causal literature (Pearl, 2009). These assumptions are required for validity of the standard IPS based estimate as well (Pearl, 2009).

Standard Causal Assumptions

Network Ignorability: $Y(\boldsymbol{z}) \perp \boldsymbol{Z} \ \forall \boldsymbol{z}$ **(A2)**

Positivity: $P(\boldsymbol{z}|\mathbf{X}) > 0 \ \forall \boldsymbol{z}$ **(A3)**

Consistency: $Y_i = Y_i(\boldsymbol{z})$ if $\boldsymbol{Z} = \boldsymbol{z}$ **(A4)**

3.2. Method

Our method relies on the insight that if Z_i are independent Bernoulli variables, then for any function $\Psi(\boldsymbol{Z})$ that admits an additive decomposition $\Psi(\boldsymbol{Z}) = \sum \Psi_i(Z_i)$, the ratio ρ_{HT} in the general HT estimator can be replaced by a lower variance $\rho_{\mathrm{RUMI}} := \sum\left(\frac{\mathbb{I}\left(Z_k|X_k\right)}{p\left(Z_k|X_k\right)} - 1\right) + 1$. Use of the weight ρ_{RUMI} in place of inverse propensity ρ_{HT} results in significantly lower variance in estimation and improved effective sample size.

This change of IPS weight can justified in the following theorem:

1. For the general mathematical statement we refer readers to Shilov and Gurevich (1978)

Proposition 1 *Under **A3**, and any function $\Psi(\boldsymbol{Z})$ admitting the decomposition $\Psi(\boldsymbol{Z}) = \sum \Psi_i(Z_i)$*

$$\mathbb{E}[\Psi(\boldsymbol{Z})\rho_{HT}] = \mathbb{E}[\Psi(\boldsymbol{Z})\rho_{RUMI}]$$

The argument behind this is similar to the reasoning in (Wen et al., 2015; Cortez-Rodriguez et al., 2023). The crux of this argument is as follows: First, since all Z are allocated independently $\mathbb{E}_p[\pi(Z_k)/p(Z_k)] = 1$ for any distribution π. Next since Ψ_k only depends on X, Z_i its contribution to the CDF is independent of the treatment at any other unit. This implies that contribution of Ψ_i in $\mathbb{E}[\rho_{HT}\Psi_i]$ is made only by the terms containing Z_i and one can ignore any contributions to the expected value from Z_{-i}. Thus $\mathbb{E}[\rho_{HT}\Psi_i] = \mathbb{E}[\pi(Z_k)/p(Z_k)\Psi_k] = \mathbb{E}[\rho_{\text{RUMI}}\Psi_i]$

Proposition 1 shows that if the target CDF F^1 is additive, then for computing it, instead of using the true IPS ρ_{HT}, one can use ρ_{RUMI}. This also provides a natural Monte-Carlo estimate by replacing the expectations with sample average i.e.

$$\begin{aligned}\hat{F}^1_{\text{RUMI}}(t) &:= \frac{1}{n}\sum_{i=1}^{n} \rho^i_{\text{RUMI}}\, \mathbb{I}\{Y_i \leqslant t\} \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Y_i \leqslant t\}\left[\sum_{j\in\mathcal{N}_i}\left(\frac{z_j}{p} - 1\right) + 1\right]\end{aligned} \tag{2}$$

Theorem 1 *Under **A1-4**, $\hat{F}^1_{RUMI}(t)$ is an unbiased pointwise consistent estimator of $F^1(t)$.*

Self-Normalization While $\hat{F}^1_{\text{RUMI}}$ can drastically reduce variance compared to $\hat{F}^1_{\text{HT}}$, it can still be subject to high variance when p is close to 0. This fact is well known in the reinforcement learning literature where such a term appears in the problem of off-policy evaluation (Precup, 2000). A common alternative there is to use the weighted (or self-normalized) IS that can further reduce the variance at the cost of bias (Tukey, 1956; Rubin, 1987; Precup, 2000). To derive such a version of $\hat{F}^1_{\text{RUMI}}$, we first note that the term z_j/p is also by itself an importance ratio (considering only Z_j as the relevant random variable) which we denote by ρ_j. Then using the same self-normalization trick we replace ρ_j by $\bar{\rho}_j = \frac{\rho_j}{\frac{1}{n}\sum_k \rho_k}$ in the formula for ρ^i_{RUMI} (Eq. (2)) to get:

$$\hat{F}^1_{\text{RUMI-WIS}}(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Y_i \leqslant t\}\left[\sum_{j\in\mathcal{N}_i}(\bar{\rho}_j - 1) + 1\right],\ \forall\ t \tag{3}$$

Theorem 2 *Under **A1-4**, $\hat{F}^1_{WIS}(\nu)$ is a pointwise asymptotically consistent estimator of $F^1(\nu)$.*

Remark 1 *The validity of the estimator does* not *require knowledge or estimation of the latent (ψ_k's) in the decomposition of the conditional CDF in Assumption 1; instead it only requires the* existence *of such functions.*

3.3. Interaction Effects

In the earlier section, we considered a density model linear in the interfering treatments (Assumption 1). However, in general, there may also exist higher order interactions from neighbouring nodes. Modeling such interactions can be important if there are strong non-linear effects such as in (Farias et al., 2022; Cortez et al., 2022). One way to model higher-order interaction structure are basing it on neighbourhood subgraphs(Aittokallio and Schwikowski, 2006; Holland and Leinhardt, 1974; Yuan et al., 2021).

Our estimator can be generalized to this setting by considering the outcome distribution to arise from a linear combination of *influences* from subsets of neighbours.

Let $\mathcal{SN}^{\beta}_i(\boldsymbol{A})$ be the set of subgraphs of size up to β that are in the neighbours of the node i. Then one can incorporate these interactions by having the cdf $F_Y(t) := \Pr(Y \leqslant t \mid \boldsymbol{Z}, X)$ to be additive in $\psi(\boldsymbol{Z}_S, t)$ where $S \in \mathcal{SN}^{\beta}_i(\boldsymbol{A})$

We present the formal assumptions and the estimator in the appendix. However we note that the variance of these estimators grows exponentially with β. Hence for many applications considering $\beta > 1$ may not give meaningful results. Furthermore there is a fundamental lower limit on the variance of unbiased non-parametric estimators (even for the average effect) as detailed by Aronow and Samii (2017) which is exponential in β and matches the asympotic complexity of our higher order estimator. Hence while we present the estimator, for practical purposes we work with the linear estimator.

3.4. Inference and Risk Metrics

The estimated CDF $\hat{F}$ can be used to compute metrics of interest as functions of the CDF (in the appendix we provide some expression for computing metrics like median, variance, CVaR etc). Many metrics are non-linear functionals of the density function, and hence can be biased even if one gets an unbiased CDF. However, since $\hat{F}$ is asymptotically consistent, values

of such metrics can be computed using the estimated CDF are also consistent.

Often practitioners would also desire to have confidence intervals to handle statistical uncertainty and verify assumptions. There is a wide literature on computing intervals for these risk functionals using the uncertainity in the distribution function F (Chandak et al., 2021b,a; Kuzborskij et al., 2020; Jiang and Huang, 2020; Chernozhukov et al., 2013). Thus one only needs to provide confidence intervals for $\hat{F}$ to get intervals for any functional of interest. Specifically, from the results in Cortez-Rodriguez et al. (2023); Shankar et al. (2024), we can provide the following claim:

Theorem 3 *Under **A1-4**, the variance of $\hat{F}_{RUMI}(t)$ is bounded above by $\frac{1}{n}F(t)(p(1-p))^{-1}d_{\mathcal{N}}^3 + \frac{F(t)(1-F(t))}{n}(p(1-p))^{-1}d_{\mathcal{N}}$ where $F(t)$ is the true underlying distribution function and $d_{\mathcal{N}}$ is the maximum degree of a node in the network graph.*

4. Experiments

The lack of ground truth counterfactuals poses a significant challenge in causal inference research, particularly when dealing with real-world data. As such researchers must rely on simulated or semi-synthetic data to precisely control the underlying data distribution and simulate counterfactual outcomes. This problem is further compounded in our case, as our focus is on the entire counterfactual distribution rather than individual measures like the mean effect.

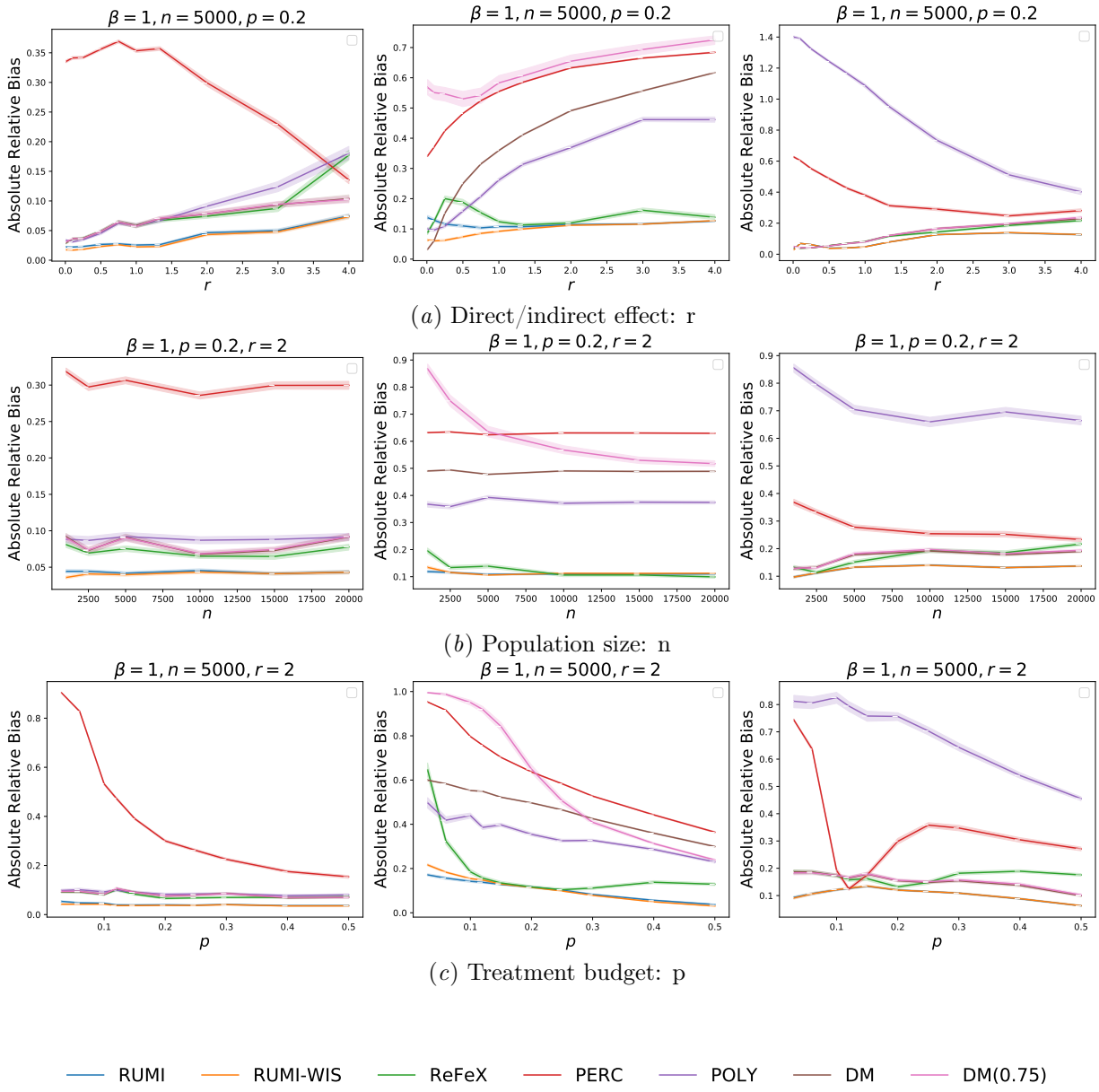
Synthetic Experiments We first experiment with interference on random graph models usually used to model social networks like the stochastic block model (SBM) and Strogatz-Watts (SW) graphs. We simulate different random SBM and SW graphs and run repeated experiments on these graphs with random treatment assignments. For these experiments, we provide an ablation study by varying the treatment probability, the strength of interference, and the size of the graphs to assess the efficacy of estimation across different ranges of parameters. For these experiments, we follow the outcome model described in Cortez-Rodriguez et al. (2023). As estimands we chose the median (equivalently the quantile at $\alpha = 0.5$) and the CVaR at $\alpha = 0.35$. For details on how these are computed from the estimated CDF we refer the readers to Appendix D.

Due to the synthetic nature of the experiments, we know the ground truth and can evaluate the bias and error of any method. We compare RUMI against some common estimators such as polynomial regression (Poly), SUTVA based estimator (DM), and other more recent models for treatment effect estimation such as ReFeX (Han and Ugander, 2023) and PERC (Zhao et al., 2022). Since the Horwitz-Thompson estimator failed to yield meaningful results, we approximate a version of the HT estimate by considering units with more than 75% of their neighbours having the same treatment as representative of a fully treated unit [2]. We draw the readers attention to the fact that of these estimators only PERC and ReFeX are capable of handling interference. Furthermore none of these estimators are designed for estimating quantiles and CVaRs. We needed to provide *additional information* about the distribution to compute these results, and was only possible as we as designers of the simulation have information about the outcomes. RUMI on the otherhand makes the linear factor assumption but is otherwise non-parametric.

In Figure 2, we illustrate the relative bias of different estimators across a range of parameters. Figure 3a plots the treatment effect estimate against the strength of interference, 2b is for varying the network size n, and 2c plots it against treatment probability p. Note that in Figure 2a, when $r = 0$, there is no interference; but even in that case only few estimator have zero bias. This is because as stated before, most estimators are only designed for average treatment effects and not distributions quantities. Furthermore, as the interference increases, baselines suffer from larger errors. We also note that a quantity like CVaR which relies on the entire CDF tail, the errors are in general larger. Finally we observe note RUMI can have high error especially when the treatment probabilities are small and graph degree is high (results in Appendix). Theorem 3 suggests that the variance of the estimator increases with low p and with high degree. Thus the error in estimation of CDF using RUMI has high variance when it comes to tails on higher degree graphs like SW model.

Health care data Next, we experiment with a real-life hospital dataset used in Song et al. (2020). The full data is not accessible, and we have used the sample data released by the authors. This sample data is not enough to relate unit covariates such as severity, age, type of condition etc. with the units. As such we limit the analysis to only two variables: the outcome vari-

2. This proxy estimator was also used in Cortez et al. (2022); Cortez-Rodriguez et al. (2023)



(a) Direct/indirect effect: r

(b) Population size: n

(c) Treatment budget: p

Figure 2: Performance of different estimators under Bernoulli design on different graph models. The charts show the variation of the model performance with : (a) Strength of interference, corresponds to the average ratio of indirect to direct effects r. (b) Population size, corresponds to the number of nodes n (c) Treatment budget, corresponds to the probability of treatment 1 (p). Y-axes represent the relative value of the absolute bias with the standard error being depicted in the shaded area. The first two columns correspond to estimating the median, while the right column corresponds to the CVaR at $\alpha = 0.35$. In each set columns, the first column corresponds to SBM and second column to SW models respectively. A failure case in CVaR estimation, due to the existence of nodes with high degree, is presented in the appendix.

able Y is the length of stay and the treatment variable Z is about off-service placement. Specifically we take $Z = 0$ if the amount of off-service placement was below the median for the data, $Z = 1$ otherwise. Furthermore lack of other covariates means that we cannot adjust for confounding and instead consider the treatments as randomized. In Fig. 3 we present the results of the Quantile Direct Effect (QDE) and Quantile Treatment Effect (QTE) estimates from RUMI over various quantiles τ [3]. While higher off-service placement has been known to increase delays (Kuntz et al., 2015), the quantile distribution is more interesting to analyze. We can see from Fig. 3 that the QTE has a peak with both the extremes having low effects. This is somewhat unusual as it suggests individuals who are mostly unaffected by the treatment. On the other hand the QDE curve is almost monotonic, suggesting. The stark difference between the qualitative nature of these curves suggest strong spillover effects. And that is in fact the case, as has been reported in Lim et al. (2024), and similar congestion based behaviour has been also noted by Kuntz et al. (2015). Specifically, off-service placement can have negligible delays over on-service treatment if the service-utilization is low i.e. there is capacity to handle the number of patients. On the other hand, shifting a patient from on-service to off-service placement can introduce major delays if the service-utilization is different. Similarly, when service-utilization is high, changing placement may not have much impact. This multi-regime behaviour may not be decipherable by only looking at aggregate treatment effect statistics and needs the estimate of the entire distribution to diagnose.

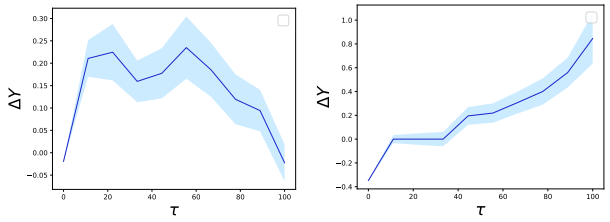


Figure 3: QTE (left) and QDE (right) for on wait times based on off-service placement across different quantiles. The U curve for total effects and the monotonic direct effect, suggest a capacity driven interference (Kuntz et al., 2015)

3. Other estimators in previous experiments cannot be applied without making further distributional assumptions

5. Conclusion

Network interference exists in many randomized control trials, and not accounting for such interference can lead to misleading conclusions. Our work focuses on estimators for the counterfactual outcomes under interference. We do so under the assumption of additive densities, which provides a structured low-variance framework for estimating counterfactual distributions in networked settings, although its validity may depend on the specific context and underlying mechanisms of interference. With both theoretical and experimental analysis, we assess the efficacy of our estimator(s) under this assumption.

Limitations and Ethical Implications While our RUMI estimator is a useful method, it's essential to acknowledge the limitations and uncertainties inherent in this method for causal inference. Additivity of the distribution is not testable without more assumptions (Han et al., 2019); and without such a test, our RUMI estimator may be unreliable. Furthermore, dealing with interference in counterfactual evaluation is a challenging problem (Chandak et al., 2021a). This area of research is still evolving, and further work is needed to develop robust methods for counterfactual distribution estimation under interference.

Our work also has potential implications for unlinked EHR records where previous treatments etc. can be considered as potential sources of interference. As such our method or its generalizations have potential connections with the responsible handling of sensitive information.

References

Alberto Abadie. Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American statistical Association*, 97 (457):284–292, 2002.

Tero Aittokallio and Benno Schwikowski. Graph-based methods for analysing networks in cell biology. *Briefings in bioinformatics*, 7(3):243–255, 2006.

Jason Altschuler, Victor-Emmanuel Brunel, and Alan Malek. Best arm identification for contaminated bandits. *Journal of Machine Learning Research*, 20 (91):1–39, 2019.

Elliott M Antman, Joseph Lau, Bruce Kupelnick, Frederick Mosteller, and Thomas C Chalmers. A comparison of results of meta-analyses of randomized control trials and recommendations of clini-

cal experts: treatments for myocardial infarction. *Jama*, 268(2):240–248, 1992.

Aronow and Samii. Estimating average causal effects under general interference. *Applied Statistics*, 2017.

Peter M Aronow and Joel A Middleton. A class of unbiased estimators of the average treatment effect in randomized experiments. *Journal of Causal Inference*, 1(1):135–154, 2013.

Susan Athey, Dean Eckles, and Guido W Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521): 230–240, 2018.

Eric Auerbach and Max Tabord-Meehan. The local approach to causal inference under network interference. *arXiv preprint arXiv:2105.03810*, 2021.

Falco J Bargagli-Stoffi, Costanza Tortù, and Laura Forastiere. Heterogeneous treatment and spillover effects under clustered network interference. *arXiv preprint arXiv:2008.00707*, 2020.

Guillaume W Basse and Edoardo M Airoldi. Model-assisted design of experiments in the presence of network-correlated outcomes. *Biometrika*, 105(4): 849–858, 2018.

Jennifer Brennan, Vahab Mirrokni, and Jean Pouget-Abadie. Cluster randomized designs for one-sided bipartite experiments. *Advances in Neural Information Processing Systems*, 35:37962–37974, 2022.

David B Brown. Large deviations bounds for estimating conditional value-at-risk. *Operations Research Letters*, 35(6):722–730, 2007.

Jing Cai, Alain De Janvry, and Elisabeth Sadoulet. Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2): 81–108, 2015.

Yash Chandak, Scott Niekum, Bruno da Silva, Erik Learned-Miller, Emma Brunskill, and Philip S Thomas. Universal off-policy evaluation. *Advances in Neural Information Processing Systems*, 34:27475–27490, 2021a.

Yash Chandak, Shiv Shankar, and Philip S Thomas. High confidence off-policy (or counterfactual) variance estimation. *Thirty-Fifth AAAI Conference on Artificial Intelligence*, 2021b.

Victor Chernozhukov and Christian Hansen. An iv model of quantile treatment effects. *Econometrica*, 73(1):245–261, 2005.

Victor Chernozhukov, Iván Fernández-Val, and Blaise Melly. Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268, 2013.

Alex Chin. Regression adjustments for estimating the global treatment effect in experiments with interference. *Journal of Causal Inference*, 7(2), 2019.

Mayleen Cortez, Matthew Eichhorn, and Christina Lee Yu. Graph agnostic estimators with staggered rollout designs under network interference. *Advances in Neural Information Processing Systems*, 2022.

Mayleen Cortez-Rodriguez, Matthew Eichhorn, and Christina Lee Yu. Exploiting neighborhood interference with low-order interactions under unit randomized design. *Journal of Causal Inference*, 11, 2023.

Iván Díaz. Efficient estimation of quantiles in missing data models. *Journal of Statistical Planning and Inference*, 190:39–51, 2017.

Miroslav Dudík, Dumitru Erhan, John Langford, and Lihong Li. Doubly robust policy evaluation and optimization. 2014.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017.

Vivek Farias, Andrew Li, Tianyi Peng, and Andrew Zheng. Markovian interference in experiments. *Advances in Neural Information Processing Systems*, 35:535–549, 2022.

Huan Gui, Ya Xu, Anmol Bhasin, and Jiawei Han. Network a/b testing: From sampling to estimation. In *Proceedings of the 24th International Conference on World Wide Web*, pages 399–409. International World Wide Web Conferences Steering Committee, 2015.

Han and Ugander. Model-based regression adjustment with model-free covariates for network interference, 2023.

Kyunghee Han, Hans-Georg Müller, and Byeong U Park. Additive functional regression for densities

as responses. *Journal of the American Statistical Association*, 2019.

Paul W Holland and Samuel Leinhardt. The statistical analysis of local structure in social networks, 1974.

Daniel G Horvitz and Donovan J Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

Michael G. Hudgens and M. Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482): 832–842, 2008.

Nan Jiang and Jiawei Huang. Minimax confidence interval for off-policy evaluation and policy optimization. *arXiv preprint arXiv:2002.02081*, 2020.

Nathan Kallus and Masatoshi Uehara. Intrinsically efficient, stable, and bounded off-policy evaluation for reinforcement learning. *Advances in neural information processing systems*, 32, 2019.

Edward H Kennedy, Sivaraman Balakrishnan, and Larry Wasserman. Semiparametric counterfactual density estimation. *arXiv preprint arXiv:2102.12034*, 2021.

Ramtin Keramati, Christoph Dann, Alex Tamkin, and Emma Brunskill. Being optimistic to be conservative: Quickly learning a cvar policy. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 4436–4443, 2020.

Kwangho Kim, Jisu Kim, and Edward H Kennedy. Causal effects based on distributional distances. *arXiv preprint arXiv:1806.02935*, 2018.

Ron Kohavi and Roger Longbotham. Online controlled experiments and a/b tests. *Encyclopedia of machine learning and data mining*, pages 1–11, 2015.

Ludwig Kuntz, Roman Mennicken, and Stefan Scholtes. Stress on the ward: Evidence of safety tipping points in hospitals. *Management Science*, 61(4):754–771, 2015.

Ilja Kuzborskij, Claire Vernade, András György, and Csaba Szepesvári. Confident off-policy evaluation and selection through self-normalized importance weighting. *arXiv preprint arXiv:2006.10460*, 2020.

James LeSage and Robert Kelley Pace. *Introduction to spatial econometrics*. Chapman and Hall/CRC, 2009.

Michael P Leung. Treatment and spillover effects under network interference. *Review of Economics and Statistics*, 102(2):368–380, 2020.

Wenrui Li, Daniel L Sussman, and Eric D Kolaczyk. Causal inference under network interference with noise. *arXiv preprint arXiv:2105.04518*, 2021.

Jong Myeong Lim, Hummy Song, and Julius J Yang. The spillover effects of capacity pooling in hospitals. *Management Science*, 2024.

Hans-Georg Müller and Fang Yao. Functional additive models. *Journal of the American Statistical Association*, 103(484):1534–1544, 2008.

Jerzy Neyman. Sur les applications de la thar des probabilities aux experiences agaricales: Essay des principle. excerpts reprinted (1990) in english. *Statistical Science*, 5(463-472):4, 1923.

Elizabeth L Ogburn, Oleg Sofrygin, Ivan Diaz, and Mark J Van der Laan. Causal inference for social network data. *arXiv preprint arXiv:1705.08527*, 2017.

Georgia Papadogeorgou, Kosuke Imai, Jason Lyall, and Fan Li. Causal inference with spatio-temporal data: Estimating the effects of airstrikes on insurgent violence in iraq. *arXiv preprint arXiv:2003.13555*, 2020.

Ben M. Parker, Steven G. Gilmour, and John Schormans. Optimal design of experiments on connected units with application to social networks. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(3):455–480, 2017. doi: https://doi.org/10.1111/rssc.12170. URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssc.12170.

Judea Pearl. *Causality*. Cambridge university press, 2009.

Jean Pouget-Abadie, Martin Saveski, Guillaume Saint-Jacques, Weitao Duan, Ya Xu, Souvik Ghosh, and Edoardo Maria Airoldi. Testing for arbitrary interference on experimentation platforms. *arXiv preprint arXiv:1704.01190*, 2017.

Jean Pouget-Abadie, Kevin Aydin, Warren Schudy, Kay Brodersen, and Vahab Mirrokni. Variance reduction in bipartite experiments through correlation clustering. *Advances in Neural Information Processing Systems*, 32, 2019.

Doina Precup. Eligibility traces for off-policy policy evaluation. *Computer Science Department Faculty Publication Series*, page 80, 2000.

James Robins and Andrea Rotnitzky. Inference for semiparametric models: Some questions and an answer - comments. *Statistica Sinica*, 11:920–936, 10 2001.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.

Donald B Rubin. The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The sir algorithm. *Journal of the American Statistical Association*, 82(398):543–546, 1987.

Fredrik Sävje, Peter M Aronow, and Michael G Hudgens. Average treatment effects in the presence of unknown interference. *The Annals of Statistics*, 49 (2):673–701, 2021.

Shiv Shankar, Ritwik Sinha, Yash Chandak, Saayan Mitra, and Madalina Fiterau. A/b testing under partial network information. In *Proceedings of the The 26th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024.

Georgiui Evgenevich Shilov and Boris Lazarevich Gurevich. *Integral, measure and derivative: a unified approach*. 1978.

Dan Siroker and Pete Koomen. *A/B testing: The most powerful way to turn clicks into customers*. John Wiley & Sons, 2015.

Hummy Song, Anita L Tucker, Ryan Graue, Sarah Moravick, and Julius J Yang. Capacity pooling in hospitals: The hidden consequences of off-service placement. *Management Science*, 66(9):3825–3842, 2020.

Daniel L Sussman and Edoardo M Airoldi. Elements of estimation theory for causal effects in the presence of network interference. *arXiv preprint arXiv:1702.03578*, 2017.

Adith Swaminathan, Akshay Krishnamurthy, Alekh Agarwal, Miro Dudik, John Langford, Damien Jose, and Imed Zitouni. Off-policy evaluation for slate recommendation. *Advances in Neural Information Processing Systems*, 30, 2017.

Panos Toulis and Edward Kao. Estimation of causal peer influence effects. In *International Conference on Machine Learning*, pages 1489–1497, 2013.

HF Tukey. Conditional monte carlo for normal samples. In *Proc. Symp. on Monte Carlo Methods*, pages 64–79. John Wiley and Sons, 1956.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337. ACM, 2013.

Davide Viviano. Experimental design under network interference. *arXiv preprint arXiv:2003.08421*, 2020.

Nikos Vlassis, Ashok Chandrashekar, Fernando Amat, and Nathan Kallus. Control variates for slate off-policy evaluation. *Advances in Neural Information Processing Systems*, 34:3667–3679, 2021.

Yu-Xiang Wang, Alekh Agarwal, and Miroslav Dudık. Optimal and adaptive off-policy evaluation in contextual bandits. In *International Conference on Machine Learning*, pages 3589–3597. PMLR, 2017.

Zheng Wen, Branislav Kveton, and Azin Ashkan. Efficient learning in large-scale combinatorial semi-bandits. In *International Conference on Machine Learning*, pages 1113–1122. PMLR, 2015.

Ted Westling and Marco Carone. A unified study of nonparametric inference for monotone functions. *Annals of statistics*, 48(2):1001, 2020.

Yuan Yuan, Kristen Altenburger, and Farshad Kooti. Causal network motifs: identifying heterogeneous spillover effects in a/b tests. In *Proceedings of the Web Conference 2021*, pages 3359–3370, 2021.

Ziyu Zhao, Kung Kuang, R Xiong, and Fei et al Wu. Learning treatment effects under heterogeneous interference in networks. *Neurips*, 2022.

Appendix A. Useful Lemmas

Lemma 2 *Suppose that $\{z_i\}_{i=1..n}$ are mutually independent, with $z_i \sim$ Bernoulli(p). Then, for any set of indices $S, S' \subset [n]$, and function f we have*

$$\mathbb{E}\Big[\prod_{i\in S}\Big(\frac{z_i}{p}-1\Big)\prod_{j\in S'} f(z_j)\Big] = \begin{cases} ((1-p)(f(1)-f(0)))^{|S\cap S'|}\mathbb{E}[f(z)]^{|S'\setminus S|} & \text{if } S\subseteq S' \\ 0 & \text{otherwise} \end{cases}$$

Proof Fix S, S'. A given index (node) i can either be only in S or only in S' or in both, with only one of the possibilities being true. Correspondingly the product, $\prod_{i\in S}\left(\frac{z_i}{p}-1\right)\prod_{j\in S'} f(z_j)$ can be factored into three exclusive products:

$$\prod_{i\in S}\Big(\frac{z_i}{p}-1\Big)\prod_{j\in S'} f(z_j) = \prod_{i\in S\setminus S'}\Big(\frac{z_i}{p}-1\Big)\prod_{k\in S\cap S'} f(z_k)\Big(\frac{z_k}{p}-1\Big)\prod_{j\in S'\setminus S} f(z_j)$$

Applying expectations and noting that z_i are mutually independent, we get:

$$\prod_{i\in S\setminus S'}\mathbb{E}[\frac{z_i}{p}-1]\prod_{k\in S\cap S'}\mathbb{E}[f(z_k)\Big(\frac{z_k}{p}-1\Big)]\prod_{j\in S'\setminus S}\mathbb{E}f(z_j) = \prod_{i\in S\setminus S'} 0 \prod_{k\in S\cap S'}\mathbb{E}[f(z_k)\Big(\frac{z_k}{p}-1\Big)]\prod_{j\in S'\setminus S}\mathbb{E}[f(z_j)]$$

The RHS can only be non zero if $S\setminus S' = \{\}$ i.e. $S\subseteq S'$.

Since $\mathbb{E}\left[f(z_k)\left(\frac{z_k}{p}-1\right)\right] = p*f(1)*(\frac{1}{p}-1)+(1-p)*f(0)*(-1) = (1-p)(f(1)-f(0))$; the RHS when it is non zero simplifies to

$$((1-p)(f(1)-f(0)))^{|S\cap S'|}\mathbb{E}[f(z)]^{|S'\setminus S|}$$

■

Corollary 3 *By putting $f(z)=z$ in Lemma 2 we get*

$$\mathbb{E}\Big[\prod_{i\in S}\Big(\frac{z_i}{p}-1\Big)\prod_{j\in S'} z_j\Big] = \begin{cases} p^{|S'\setminus S|}(1-p)^{|S|} & \text{if } S\subseteq S' \\ 0 & \text{otherwise} \end{cases}$$

Lemma 4 *For any sets $S', \mathcal{N}_i$ such that $S'\subseteq\mathcal{N}_i$*

$$\mathbb{E}\left[\prod_{k\in S'}(z_k)\sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big(\prod_{j\in S}\frac{z_j-p}{p}\Big)\right] = 1$$

Proof

Applying Theorem 3 with $f_i(z)=z_i$ we get:

$$\begin{aligned}\mathbb{E}\left[\prod_{k\in S'}(z_k)\sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big(\prod_{j\in S}\frac{z_j-p}{p}\Big)\right] &= \sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big[p^{|S'/S]}(1-p)^{|S'\cap S|}\mathbb{I}[S\subseteq S']\Big] \\ &\overset{(b)}{=} \sum_{\substack{S\subseteq S'\\|S|\leqslant\beta}}\Big[p^{|S'/S]}(1-p)^{|S'\cap S|}\Big] \qquad \text{(S1)}\end{aligned}$$

(b) follows from that $\mathbb{I}[S \subseteq S']$ will filter any non subset of S'

$$= \sum_{\substack{S \subseteq S' \\ |S| \leqslant \beta}} p^{|S'|} \left[p^{-|S|} (1-p)^{|S|} \right]$$

$$= \sum_{\substack{S \subseteq S' \\ |S| \leqslant \beta}} p^{|S'|} \left[(\frac{1}{p} - 1)^{|S|} \right] \tag{S2}$$

Note that the terms in the sum S2 only depend on sizes of the subset S and not the elements in it. Hence the sum S1 can be rewritten as:

$$= p^{|S'|} \sum_{\substack{r \leqslant \beta \\ r \leqslant |S'|}} \binom{|S'|}{r} \left[(\frac{1}{p} - 1)^{r} \right]$$

$$= p^{|S'|} \left[\sum_{\substack{r \leqslant \beta \\ r \leqslant |S'|}} \binom{|S'|}{r} (\frac{1}{p} - 1)^{r} \right] \tag{S3}$$

If $|S'| \leqslant \beta$, we are summing over all subsets of S', and the constraint of $r \leqslant \beta$ is redundant. Then by applying binomial theorem we get.

$$\overset{(c)}{=} p^{|S'|} (1 + (\frac{1}{p} - 1))^{|S'|} = 1$$

where in (c) we used binomial theorem as $\sum_r \binom{n}{r} x^r = (1+x)^n$ ■

Appendix B. Interaction/Motif Model Estimator

Motif Additive Density

Assumption 2 *The (conditional) cumulative density function (CDF) of the outcomes can be written as a linear combination of motif latent functions:*

$$F_Y(t) = \sum_{S \in \mathcal{SN}_i^{\beta}(\boldsymbol{A})} \psi_S(Z_S, t), \ \forall t \tag{A5}$$

where $F_Y(t) := \Pr(Y \leqslant t \mid \boldsymbol{Z}, X)$.

Note that, if we consider only dyads, i.e. we set $\beta = 1$ and consider all such dyadic elements in $\mathcal{N}_i$ then **A5** is equivalent to **A1**. Considering $\beta > 1$ allows non-linear interaction between the nodes.

Estimator We have

$$ {}_{\beta}\hat{F}^1_{\text{RUMI}}(t) \coloneqq \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{Y_i \leqslant t\} \sum_{\substack{S \subseteq \mathcal{N}_i \\ |S| \leqslant \beta}} \Big(\prod_{j \in S} \frac{z_j - p}{p} \Big) \tag{4}$$

where by convention a product over terms in an empty set evaluates to 1.

Remark 5 *Note that the estimator* ${}_{\beta}\hat{F}^1_{RUMI}$ *devolves to* $\hat{F}^1_{RUMI}$ *when we set* $\beta = 1$. *The additional 1 in Eq. (2) comes from the contribution corresponding to* $S = \phi = \{\}$ *in Eq. (4)*

Theorem 4 *Under* ***A2-5****,* ${}_{\beta}\hat{F}^1_{RUMI}$ *is an unbiased pointwise consistent estimator of* $F^1(\nu)$.

Self-Normalized Version: Once again, expressing ${}_\beta\hat{F}^1_{\text{RUMI}}$ using $\rho_i = z_j/p$ and replacing ρ by $\bar{\rho}$ we can get a self-normalized version given by:

$$ {}_\beta\hat{F}^1_{\text{RUMI-WIS}}(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Y_i \leqslant t\} \sum_{\substack{S\subseteq\mathcal{M}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}(\bar{\rho}_j - 1)\Big) $$

Appendix C. Proofs

Proof of Theorem 4

Proof First note that for any bernoulli variable Z, any positive power Z^n of Z is the same bernoulli variable as Z. Hence the latent density function $\psi_{S',\boldsymbol{Z}_{S'},t}$ in the motif model of size β can be written as a product of Z_j for $j \in S'$

$$ \psi_{S',\boldsymbol{Z}_{S'},t} = c_{i,S',t} \prod_{j\in S'} \mathbb{I}[z_j = 1] \tag{5} $$

Lets fix a real number t. Next consider the variable $Y_i < t$. This is a bernoulli random variable with the probability of success given by $Pr(Y_i < t)$ which by definition is $F_Y(t)$ (the cdf of Y). Hence the indicator random variable $\mathbb{I}[Y_i < t]$ can be written as

$$ \mathbb{I}[Y_i < t] = F(t) + \epsilon $$

where ϵ is a zero mean random variable. Note that while ϵ is dependent on $F(t)$, given the treatment allocation $\boldsymbol{Z}$ it acts as an independent variable. Specifically, for any integrable function g of $\boldsymbol{Z}$ we have

$$ \mathbb{E}[\epsilon g(Z)|\boldsymbol{Z}] = \mathbb{E}[\epsilon|\boldsymbol{Z}]\mathbb{E}[g(Z)|\boldsymbol{Z}] = 0 $$

Now lets consider the expected value of the estimator ${}_\beta\hat{F}^1_{\text{RUMI}}$. We have

$$ \mathbb{E}[{}_\beta\hat{F}^1_{\text{RUMI}}](t) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\{Y_i \leqslant t\} \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] \tag{6} $$

$$ = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} (F(t) + \epsilon_i) \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] \tag{7} $$

$$ = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} F(t) \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] + \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \epsilon_i \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] \tag{8} $$

$$ = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[F(t) \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] + \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\epsilon_i \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] \tag{9} $$

$$ \overset{a}{=} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\sum \sum_{\substack{S'\subseteq\mathcal{N}_i \\ |S'|\leqslant\beta}} \psi_{S',\boldsymbol{Z}_{S'},t} \sum_{\substack{S\subseteq\mathcal{N}_i \\ |S|\leqslant\beta}} \Big(\prod_{j\in S}\frac{z_j - p}{p}\Big)\right] + 0 \tag{10} $$

where in (a) we have replaced $F(t)$ with our assumption A5, and used the previous results that $\mathbb{E}[\epsilon g(Z)] = 0$

$$\stackrel{b}{=} \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\sum_{\substack{S'\subseteq\mathcal{N}_i\\|S'|\leqslant\beta}} c_{i,S',t}\prod_{j\in S'}\mathbb{I}[z_j=1]\sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big(\prod_{j\in S}\frac{z_j-p}{p}\Big)\right] \tag{11}$$

where in (b) we used the motif representation of the latent functions in Eq. (5)

$$= \mathbb{E}\left[\frac{1}{n}\sum_i\sum_{\substack{S'\subseteq\mathcal{N}_i\\|S'|\leqslant\beta}} c_{i,S'}\prod_{j\in S'}\mathbb{I}[z_j=1]\sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big(\prod_{j\in S}\frac{z_j-p}{p}\Big)\right] \tag{12}$$

$$= \mathbb{E}\left[\frac{1}{n}\sum_i\sum_{\substack{S'\subseteq\mathcal{N}_i\\|S'|\leqslant\beta}} c_{i,S'}\prod_{j\in S'} z_j\sum_{\substack{S\subseteq\mathcal{N}_i\\|S|\leqslant\beta}}\Big(\prod_{j\in S}\frac{z_j-p}{p}\Big)\right] \tag{13}$$

$$\stackrel{c}{=} \left[\frac{1}{n}\sum_i\sum_{\substack{S'\subseteq\mathcal{N}_i\\|S'|\leqslant\beta}} c_{i,S'}1\right] \tag{14}$$

where we have used in (c) Theorem 4

$$= \frac{1}{n}\sum_i\sum_{\substack{S'\subseteq\mathcal{N}_i\\|S'|\leqslant\beta}} c_{i,S'} \tag{15}$$

$$\stackrel{d}{=} \frac{1}{n}\sum_i\sum_{S'\subset\mathcal{N}_i}\psi_{S'}(Z_{S'}=\mathbf{1},t) \tag{16}$$

where we have used in (d) the motif representation of the latent functions in Eq. (5)

$$= \frac{1}{n}\sum_i F^1(t) = F^1(t) \tag{17}$$

■

Proof of Theorem 1

Proof Theorem 1 can be seen to derive from Theorem 4 by placing $\beta = 1$ ■

Proof of Theorem 2

Proof Note that as all z_i are independent, so are all $\rho_i = z_i/p$. As such we can apply Kolmogorove's strong law of large numbers to any linear combination of these variables and get the following

$$\lim_{n\to\infty}\frac{1}{n}\sum_j\sum_{k\in\mathcal{N}_j}\frac{1}{|\mathcal{N}_k|}\rho_k \xrightarrow{a.s.} \sum_j\sum_{k\in\mathcal{N}_j}\frac{1}{|\mathcal{N}'_k|}\mathbb{E}[\rho_k] = 1$$

Since $1/t$ is a continuous function at $t = 1$, we can apply Slutsky continuous mapping theorem on the definition of $\bar{\rho}$ to get that

$$\text{as}\lim_{n\to\infty}: \quad \bar{\rho}_i \to \rho_i.$$

This along with Theorem 4 proves the asymptotic consistency of ${}_\beta\hat{F}^1_{\text{RUMI-WIS}}$. Since $\hat{F}^1_{\text{RUMI-WIS}}$ is the restriction of ${}_\beta\hat{F}^1_{\text{RUMI-WIS}}$ to $\beta = 1$, we automatically get the consistency of $\hat{F}^1_{\text{RUMI-WIS}}$ as well. ■

Proof of Theorem 3

Proof

We first state a useful result from Section 5.3 in Cortez-Rodriguez et al. (2023) For the estimator $\hat{\tau}^\beta = \frac{1}{n}\sum_{i=1}^n Y_i \sum_{\substack{S\subseteq\mathcal{M}_i \\ |S|\leqslant\beta}} \left(\prod_{j\in S} \frac{z_j - p}{p}\right)$, we have the following bound:

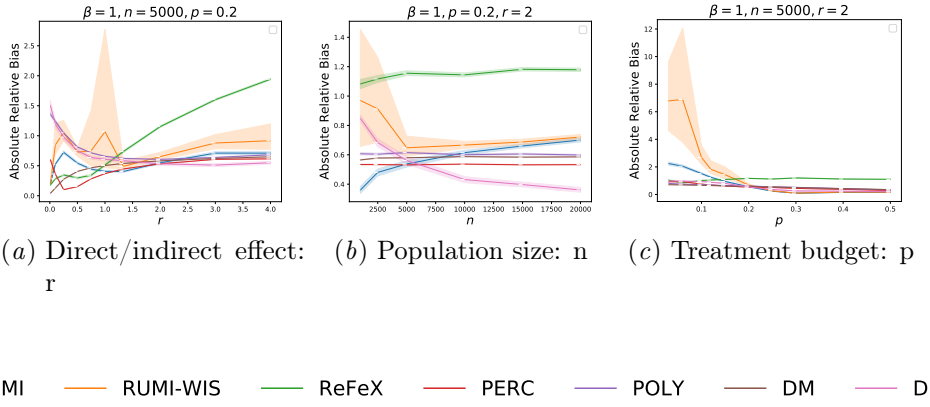
$$\text{Var}[\hat{\tau}^\beta] \leqslant \frac{1}{n} Y_{\max}^2 (p(1-p))^{-\beta} d_{\mathcal{N}}^{2\beta+2} + \frac{\sigma^2}{n} d_{\mathcal{M}}^{2\beta} (p(1-p))^{-\beta}$$

here Y_i is a stochastic polynomial of order β, where $Y_{\max}$ is an upper bound for all Y_i and σ^2 is the variance of Y_i.

Next we note that the estimator ${}_\beta\hat{F}^1_{\text{RUMI}}$ is of the same form as required in the previous result. Here the outcome polynomial Y_i are replaced by the true cdf value $F(t)$, which by the nature of being a cdf is upper-bounded by 1. Secondly the variance of ϵ can be seen from the classic result on variance of Bernoulli variables. Putting in the values of $Y_{\max}$ and σ^2 in this theorem of Cortez-Rodriguez et al. (2023), gives the required result. ■

Appendix D. Additional Results

Here we present some additional results for CVaR estimation. We can see that CVaR tends to be more biased than quantiles. Furthermore when the degree of the nodes is high, the variance of the estimator can be quite high as can be seen from Figure 4.



(a) Direct/indirect effect: r (b) Population size: n (c) Treatment budget: p

Figure 4: Performance of different estimators under Bernoulli design for CVaR estimation on SW graphs. The charts show the variation of the model performance with : (a) Strength of interference, corresponds to the average ratio of indirect to direct effects r. (b) Population size, corresponds to the number of nodes n (c) Treatment budget, corresponds to the probability of treatment 1 (p).

Appendix E. Computing Risk Metrics

Using an estimate $\hat{F}$ of the true distribution F, any functional $\Xi(F_\pi)$ of the distribution can now be estimated using $\Xi(\hat{F}_n)$.

However some parameters like the variance, quantiles and tail-risk are typically defined using the inverse CDF $F_\pi^{-1}(\alpha)$ and the density function $\mathrm{d}F$ instead of the cumulative distribution F. To compute such metrics, we utilize the method described in Chandak et al. (2021a) and Brown (2007), which relies on order statistics of the observed data.

We first define some some terms:

- $F_\pi^{-1}(\alpha)$ represents the α-quantile of the distribution function F.
- $\mathrm{d}F$ represents the probability density function of the distribution F.
- $Y_{(i)}$ denotes the ith order statistic of the observed outcomes, where $Y_{(1)}$ is the smallest observed value and $Y_{(n)}$ is the largest observed value, assuming n observations.
- σ_F^2 represents the variance
- $\mu(F)$ represents the mean
- Q_F^α stands for the α-quantile
- CVaR_F^α is the conditional value at risk (also known as tail-risk) is a tail-sensitive measure) at α probability

We can create estimators of the inverse CDF and the probability density as,

$$\hat{F}^{-1}(\alpha) \coloneqq \min\left\{y \in (Y_{(i)})_{i=1}^n \middle| \hat{F}(y) \geqslant \alpha\right\}, \qquad \mathrm{d}\hat{F}(Y_{(i)}) \coloneqq \hat{F}(Y_{(i)}) - \hat{F}(Y_{(i-1)}), \tag{18}$$

where $\mathrm{d}\hat{F}(\nu) \coloneqq 0$ if $\nu \neq Y_{(i)}$ for any $i \in (1, \ldots, n)$. Note that this is a discrete approximation to the density. We use the $\hat{}$ notation to distinguish the true values versus their estimates. Using (18), we can define estimators for variance, quantiles, inter-quantile range (IQR) and CVaR as:

$$\sigma_{\hat{F}}^2 \coloneqq \sum_{i=1}^n \mathrm{d}\hat{F}(Y_{(i)})\left(Y_{(i)} - \mu(\hat{F})\right)^2 \qquad Q_{\hat{F}}^\alpha \coloneqq \hat{F}^{-1}(\alpha) \tag{19}$$

$$\mathrm{IQR}_{\hat{F}}^{\alpha_1,\alpha_2} \coloneqq Q_{\hat{F}}^{\alpha_2} - Q_{\hat{F}}^{\alpha_1}, \qquad \mathrm{CVaR}_{\hat{F}}^\alpha \coloneqq \frac{1}{\alpha}\sum_{i=1}^n \mathrm{d}\hat{F}(Y_{(i)}) Y_{(i)} \mathbb{I}\left[Y_{(i)} \leqslant Q_{\hat{F}}^\alpha\right]. \tag{20}$$

Notice that all of these are nonlinear in F and $\hat{F}$, therefore, even if we have an unbiased CDF estimate $\hat{F}$; the sample estimates of a functional $\Xi(\hat{F})$, can be a biased. However, the consistency of $\hat{F}$, and the fact that all the above metrics are continuous in $\hat{F}$ (except CVar) means that these estimates are also consistent.